Einar Stole 1814 11/14
305-4507

CRF Errors Corrected by the STIC Systems Branch

Serial Number: 08/898,560

CRF Processing Date: 11/21/99
Edited by: _____
Verified by: _____ (STIC staff)

☐ Changed a file from non-ASCII to ASCII

☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.

☐ Edited a format error in the Current Application Data section, specifically: _____

ENTERED

☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____.

☐ Added the mandatory heading and subheadings for "Current Application Data".

☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.

☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: _____

☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: _____

☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: _____

☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.

☐ Inserted colons after headings/subheadings. Headings edited included: _____

☐ Deleted extra, invalid, headings used by an applicant, specifically: _____

☐ Deleted: ☐ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as _____.

☐ Inserted mandatory headings, specifically: _____

☐ Corrected an obvious error in the response, specifically: _____

☐ Edited identifiers where upper case is used but lower case is required, or vice versa.

☐ Corrected an error in the Number of Sequences field, specifically: _____

☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.

☐ Deleted ***ending*** stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: _____

☑ Other: Seqs 1, 2 – changed "KEY" to "NAME/KEY" under (ix) FEATURE:

***Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.**

3/1/95

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

INPUT SET: S21525.raw

This Raw Listing contains the General Information Section and up to the first 5 pages.

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Hiroyuki NAKANE, Chikara OHTO, Shinichi OHNUMA,
Kazutake HIROOKA, Tokuzo NISHINO

(ii) TITLE OF INVENTION: Farnesyl Diphosphate Synthase

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Kenyon & Kenyon
(B) STREET: One Broadway
(C) CITY: New York
(D) STATE: NY
(E) COUNTRY: USA
(F) ZIP: 10004

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 3" Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
(D) SOFTWARE: WordPerfect 6.1 Windows

(vi) CURRENT APPLICATION DATA:
--> **(A) APPLICATION NUMBER: Not Yet Issued**
(B) FILING DATE: Concurrent Herewith
--> **(C) CLASSIFICATION: Not Yet Issued**

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: JP 8-213211
(B) FILING DATE: 24-JUL-96

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Edward W. Greason
(B) REGISTRATION NUMBER: 18,918
(C) REFERENCE/DOCKET NUMBER: 77670/495

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212)425-7200
(B) TELEFAX: (212)425-5288

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 330 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Sulfolobus acidocaldarius
(B) STRAIN: ATCC 33909

(ix) FEATURE:
(A) NAME/KEY: Asp-rich domain
(B) LOCATION: 82-86

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn
                  5                  10                  15
Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu
             20                  25                  30
Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu
         35                  40                  45
Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala
     50                  55                  60
Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val
 65                  70                  75                  80
His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr
                 85                  90                  95
Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu
            100                 105                 110
Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu
        115                 120                 125
Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile
    130                 135                 140
Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg
145                 150                 155                 160
Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr
                165                 170                 175
Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly
            180                 185                 190
Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu
        195                 200                 205
Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp
    210                 215                 220
Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
225                 230                 235                 240

Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu
                245                 250                 255
Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu
            260                 265                 270
Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
        275                 280                 285
```

Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
290 295 300
Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
305 310 315 320
Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
325 330

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 993 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Sulfolobus acidocaldarius
(B) STRAIN: ATCC 33909

(ix) FEATURE:
(A) NAME/KEY: Asp-rich domain coding
(B) LOCATION: 246-258

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGAGTTACT TTGACAACTA TTTTAATGAG ATTGTTAATT CTGTAAACGA CATTATTAAG 60

AGCTATATAT CTGGAGATGT TCCTAAACTA TATGAAGCCT CATATCATTT GTTTACATCT 120

GGAGGTAAGA GGTTAAGACC ATTAATCTTA ACTATATCAT CAGATTTATT CGGAGGACAG 180

AGAGAAAGAG CTTATTATGC AGGTGCAGCT ATTGAAGTTC TTCATACTTT TACGCTTGTG 240

CATGATGATA TTATGGATCA AGATAATATC AGAAGAGGGT TACCCACAGT CCACGTGAAA 300

TACGGCTTAC CCTTAGCAAT ATTAGCTGGG GATTTACTAC ATGCAAAGGC TTTTCAGCTC 360

TTAACCCAGG CTCTTAGAGG TTTGCCAAGT GAAACCATAA TTAAGGCTTT CGATATTTTC 420

ACTCGTTCAA TAATAATTAT ATCCGAAGGA CAGGCAGTAG ATATGGAATT TGAGGACAGA 480

ATTGATATAA AGGAGCAGGA ATACCTTGAC ATGATCTCAC GTAAGACAGC TGCATTATTC 540

TCGGCATCCT CAAGTATAGG CGCACTTATT GCTGGTGCTA ATGATAATGA TGTAAGACTG 600

ATGTCTGATT TCGGTACGAA TCTAGGTATT GCATTTCAGA TTGTTGACGA TATCTTAGGT 660

CTAACAGCAG ACGAAAAGGA ACTTGGAAAG CCTGTTTTTA GTGATATTAG GGAGGGTAAA 720

AAGACTATAC TTGTAATAAA AACACTGGAG CTTTGTAAAG AGGACGAGAA GAAGATTGTC 780

CTAAAGGCGT TAGGTAATAA GTCAGCCTCA AAAGAAGAAT TAATGAGCTC AGCAGATATA 840

ATTAAGAAAT ACTCTTTAGA TTATGCATAC AATTTAGCAG AGAAATATTA TAAAAATGCT 900

ATAGACTCTT TAAATCAAGT CTCCTCTAAG AGTGATATAC CTGGAAAGGC TTTAAAATAT 960

CTAGCTGAAT TTACGATAAG AAGGAGAAAA TAA 993

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CATACTTTTT TCCTTGTGGC TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CATACTTTTT TCCTTGTGCT TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CATACTTATT TCCTTGTGCT TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH:37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CATACTTATT TCCTTGTGGC TGATGATATC ATGGATC 37

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTTCTTCATA CTTATTCGCT TATTCATGAT AGTATT 36

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATTCATGATG ATCTTCCATC GATGGATCAA GAT 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid

****** PREVIOUSLY ERRORED SEQUENCES - EDITED *****

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Sulfolobus acidocaldarius
(B) STRAIN: ATCC 33909

(ix) FEATURE:
(A) NAME/KEY: Asp-rich domain
(B) LOCATION: 82-86

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn
                5                  10                  15
Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu
            20                  25                  30
Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu
        35                  40                  45
Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala
    50                  55                  60
Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val
65                  70                  75                  80
His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr
                85                  90                  95
Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu
            100                 105                 110
Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu
        115                 120                 125
Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile
    130                 135                 140
Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg
145                 150                 155                 160
Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr
                165                 170                 175
Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly
            180                 185                 190
Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu
        195                 200                 205
Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp
    210                 215                 220
```

```
 91  Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
 92  225                 230                 235                 240
 93
 94  Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu
 95                  245                 250                 255
 96  Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu
 97              260                 265                 270
 98  Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
 99          275                 280                 285
100  Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
101      290                 295                 300
102  Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
103  305                 310                 315                 320
104  Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
105                  325                 330
106
```

```
107     (2) INFORMATION FOR SEQ ID NO:2:
108
109  (i) SEQUENCE CHARACTERISTICS:
110     (A) LENGTH: 993 base pairs
111     (B) TYPE: nucleic acid
112     (C) STRANDEDNESS: double
113     (D) TOPOLOGY: linear
114
115  (ii) MOLECULE TYPE: genomic DNA
116
117  (vi) ORIGINAL SOURCE:
118     (A) ORGANISM: Sulfolobus acidocaldarius
119     (B) STRAIN: ATCC 33909
120
121  (ix) FEATURE:
122     (A) NAME/KEY: Asp-rich domain coding
123     (B) LOCATION: 246-258
124
125  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
126
127  ATGAGTTACT TTGACAACTA TTTTAATGAG ATTGTTAATT CTGTAAACGA CATTATTAAG   60
128
129  AGCTATATAT CTGGAGATGT TCCTAAACTA TATGAAGCCT CATATCATTT GTTTACATCT  120
130
131  GGAGGTAAGA GGTTAAGACC ATTAATCTTA ACTATATCAT CAGATTTATT CGGAGGACAG  180
132
133  AGAGAAAGAG CTTATTATGC AGGTGCAGCT ATTGAAGTTC TTCATACTTT TACGCTTGTG  240
134
135  CATGATGATA TTATGGATCA AGATAATATC AGAAGAGGGT TACCCACAGT CCACGTGAAA  300
136
137  TACGGCTTAC CCTTAGCAAT ATTAGCTGGG GATTTACTAC ATGCAAAGGC TTTTCAGCTC  360
138
139  TTAACCCAGG CTCTTAGAGG TTTGCCAAGT GAAACCATAA TTAAGGCTTT CGATATTTTC  420
140
141  ACTCGTTCAA TAATAATTAT ATCCGAAGGA CAGGCAGTAG ATATGGAATT TGAGGACAGA  480
142
```

```
143   ATTGATATAA AGGAGCAGGA ATACCTTGAC ATGATCTCAC GTAAGACAGC TGCATTATTC  540
144
145   TCGGCATCCT CAAGTATAGG CGCACTTATT GCTGGTGCTA ATGATAATGA TGTAAGACTG  600
146
147   ATGTCTGATT TCGGTACGAA TCTAGGTATT GCATTTCAGA TTGTTGACGA TATCTTAGGT  660
148
149   CTAACAGCAG ACGAAAAGGA ACTTGGAAAG CCTGTTTTTA GTGATATTAG GGAGGGTAAA  720
150
151   AAGACTATAC TTGTAATAAA AACACTGGAG CTTTGTAAAG AGGACGAGAA GAAGATTGTC  780
152
153   CTAAAGGCGT TAGGTAATAA GTCAGCCTCA AAAGAAGAAT TAATGAGCTC AGCAGATATA  840
154
155   ATTAAGAAAT ACTCTTTAGA TTATGCATAC AATTTAGCAG AGAAATATTA TAAAAATGCT  900
156
157   ATAGACTCTT TAAATCAAGT CTCCTCTAAG AGTGATATAC CTGGAAAGGC TTTAAAATAT  960
158
159   CTAGCTGAAT TTACGATAAG AAGGAGAAAA TAA                               993
160
161
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION ***US/08/898,560***

DATE: 11/12/97
TIME: 21:24:41

INPUT SET: S21525.raw

Line	Error	Original Text
27	Wrong application Serial Number	(A) APPLICATION NUMBER: Not Yet Issued
29	Wrong Classification	(C) CLASSIFICATION: Not Yet Issued

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

INPUT SET: S21525.raw

This Raw Listing contains the General Information Section and those Sequences containing ERRORS.

Does Not Comply
Corrected Diskette Needed

```
SEQUENCE LISTING

(1)    General Information:

(i) APPLICANTS: Hiroyuki NAKANE,  Chikara OHTO,  Shinichi OHNUMA,
Kazutake HIROOKA,  Tokuzo NISHINO

(ii) TITLE OF INVENTION: Farnesyl Diphosphate Synthase

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:
   (A) ADDRESSEE: Kenyon & Kenyon
   (B) STREET: One Broadway
   (C) CITY: New York
   (D) STATE: NY
   (E) COUNTRY: USA
   (F) ZIP: 10004

(v) COMPUTER READABLE FORM:
   (A) MEDIUM TYPE: 3" Floppy disk
   (B) COMPUTER: IBM PC compatible
   (C) OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
   (D) SOFTWARE: WordPerfect 6.1 Windows

(vi) CURRENT APPLICATION DATA:
--> (A) APPLICATION NUMBER: Not Yet Issued
   (B) FILING DATE: Concurrent Herewith
--> (C) CLASSIFICATION: Not Yet Issued

(vii) PRIOR APPLICATION DATA:
   (A) APPLICATION NUMBER: JP 8-213211
   (B) FILING DATE: 24-JUL-96

(viii) ATTORNEY/AGENT INFORMATION:
   (A) NAME: Edward W. Greason
   (B) REGISTRATION NUMBER: 18,918
   (C) REFERENCE/DOCKET NUMBER: 77670/495

(ix) TELECOMMUNICATION INFORMATION:
   (A) TELEPHONE: (212)425-7200
   (B) TELEFAX: (212)425-5288
```

ERRORED SEQUENCES FOLLOW:

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/898,560*

INPUT SET: S21525.raw

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Sulfolobus acidocaldarius
(B) STRAIN: ATCC 33909

(ix) FEATURE: → NAME/KEY:
--> **(A) KEY: Asp-rich domain**
(B) LOCATION: 82-86

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

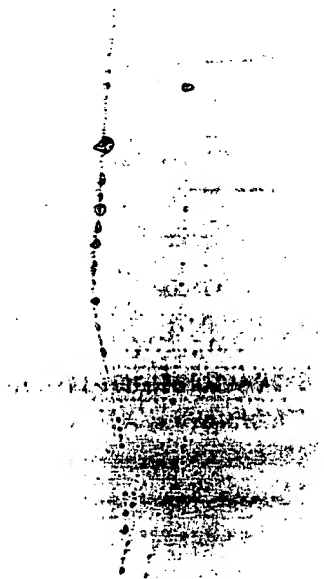
```
Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val Asn
                5                  10                  15
Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu Tyr Glu
            20                  25                  30
Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu Arg Pro Leu
        35                  40                  45
Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln Arg Glu Arg Ala
    50                  55                  60
Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His Thr Phe Thr Leu Val
65                  70                  75                  80
His Asp Asp Ile Met Asp Gln Asp Asn Ile Arg Arg Gly Leu Pro Thr
                85                  90                  95
Val His Val Lys Tyr Gly Leu Pro Leu Ala Ile Leu Ala Gly Asp Leu
            100                 105                 110
Leu His Ala Lys Ala Phe Gln Leu Leu Thr Gln Ala Leu Arg Gly Leu
        115                 120                 125
Pro Ser Glu Thr Ile Ile Lys Ala Phe Asp Ile Phe Thr Arg Ser Ile
    130                 135                 140
Ile Ile Ile Ser Glu Gly Gln Ala Val Asp Met Glu Phe Glu Asp Arg
145                 150                 155                 160
Ile Asp Ile Lys Glu Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr
                165                 170                 175
Ala Ala Leu Phe Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly
            180                 185                 190
Ala Asn Asp Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu
        195                 200                 205
Gly Ile Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp
    210                 215                 220
Glu Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
225                 230                 235                 240

Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp Glu
                245                 250                 255
Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser Lys Glu
```

```
 97                 260                 265                 270
 98  Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser Leu Asp Tyr
 99          275                 280                 285
100  Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Ser Leu
101      290                 295                 300
102  Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly Lys Ala Leu Lys Tyr
103  305                 310                 315                 320
104  Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
105                  325                 330
106
```

```
107     (2) INFORMATION FOR SEQ ID NO:2:
108
109  (i) SEQUENCE CHARACTERISTICS:
110     (A) LENGTH: 993 base pairs
111     (B) TYPE: nucleic acid
112     (C) STRANDEDNESS: double
113     (D) TOPOLOGY: linear
114
115  (ii) MOLECULE TYPE: genomic DNA
116
117  (vi) ORIGINAL SOURCE:
118     (A) ORGANISM: Sulfolobus acidocaldarius
119     (B) STRAIN: ATCC 33909
120
121  (ix) FEATURE:
122     (A) KEY: Asp-rich domain coding
123     (B) LOCATION: 246-258
124
125  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
126
127  ATGAGTTACT TTGACAACTA TTTTAATGAG ATTGTTAATT CTGTAAACGA CATTATTAAG   60
128
129  AGCTATATAT CTGGAGATGT TCCTAAACTA TATGAAGCCT CATATCATTT GTTTACATCT  120
130
131  GGAGGTAAGA GGTTAAGACC ATTAATCTTA ACTATATCAT CAGATTTATT CGGAGGACAG  180
132
133  AGAGAAAGAG CTTATTATGC AGGTGCAGCT ATTGAAGTTC TTCATACTTT TACGCTTGTG  240
134
135  CATGATGATA TTATGGATCA AGATAATATC AGAAGAGGGT TACCCACAGT CCACGTGAAA  300
136
137  TACGGCTTAC CCTTAGCAAT ATTAGCTGGG GATTTACTAC ATGCAAAGGC TTTTCAGCTC  360
138
139  TTAACCCAGG CTCTTAGAGG TTTGCCAAGT GAAACCATAA TTAAGGCTTT CGATATTTTC  420
140
141  ACTCGTTCAA TAATAATTAT ATCCGAAGGA CAGGCAGTAG ATATGGAATT TGAGGACAGA  480
142
143  ATTGATATAA AGGAGCAGGA ATACCTTGAC ATGATCTCAC GTAAGACAGC TGCATTATTC  540
144
145  TCGGCATCCT CAAGTATAGG CGCACTTATT GCTGGTGCTA ATGATAATGA TGTAAGACTG  600
146
147  ATGTCTGATT TCGGTACGAA TCTAGGTATT GCATTTCAGA TTGTTGACGA TATCTTAGGT  660
148
```

```
149  CTAACAGCAG ACGAAAAGGA ACTTGGAAAG CCTGTTTTTA GTGATATTAG GGAGGGTAAA  720
150
151  AAGACTATAC TTGTAATAAA AACACTGGAG CTTTGTAAAG AGGACGAGAA GAAGATTGTC  780
152
153  CTAAAGGCGT TAGGTAATAA GTCAGCCTCA AAAGAAGAAT TAATGAGCTC AGCAGATATA  840
154
155  ATTAAGAAAT ACTCTTTAGA TTATGCATAC AATTTAGCAG AGAAATATTA TAAAAATGCT  900
156
157  ATAGACTCTT TAAATCAAGT CTCCTCTAAG AGTGATATAC CTGGAAAGGC TTTAAAATAT  960
158
159  CTAGCTGAAT TTACGATAAG AAGGAGAAAA TAA                               993
160
161
```



SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/898,560*

DATE: 11/12/97
TIME: 16:24:35

INPUT SET: S21525.raw

Line	Error	Original Text
27	Wrong application Serial Number	(A) APPLICATION NUMBER: Not Yet Issued
29	Wrong Classification	(C) CLASSIFICATION: Not Yet Issued
58	Unknown or Misplaced Identifier	(A) KEY: Asp-rich domain
122	Unknown or Misplaced Identifier	(A) KEY: Asp-rich domain coding